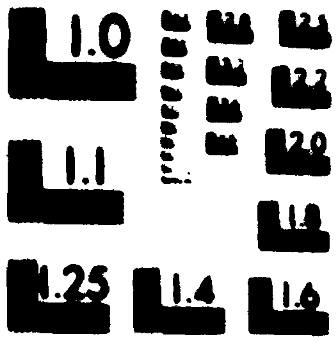
AD-A181 603 MAXIMUM LIKELIHOOD ESTIMATION OF A CLASS OF NON-GAUSSIAN DENSITIES WITH A.. (U) RICE UNIV HOUSTON TX DEPT OF ELECTRICAL AND COMPUTER ENGINEER.. 1/1

UNCLASSIFIED T T PHAM ET AL. 1987 N00014-85-K-0152 F/G 12/2 NL


1.0
1.1
1.25
1.4
1.6
1.8
2.0
2.2
2.5

(12)

FILE COPY

RICE UNIVERSITY

GEORGE R. BROWN SCHOOL OF ENGINEERING

AD-A181 683

MAXIMUM LIKELIHOOD ESTIMATION OF A CLASS OF NON-GAUSSIAN DENSITIES WITH APPLICATION TO DECONVOLUTION

Trung T. Pham and Rui J. P. de Figueiredo

Department of Electrical and Computer Engineering
Rice University
Houston, TX 77251-1892

Proceedings of the [illegible] ICASSP Dallas Texas

DTIC
ELECTE
JUN 05 1987
S D
D

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

HOUSTON, TEXAS

87 5 20 121

(12)

MAXIMUM LIKELIHOOD ESTIMATION OF A CLASS OF NON-GAUSSIAN DENSITIES WITH APPLICATION TO DECONVOLUTION

Trung T. Pham and Rui J.P. de Figueiredo

Department of Electrical and Computer Engineering
Rice University
Houston, TX 77251-1892

Proceedings of the 1987 ICASSP Dallas Texas

[illegible]

Truong T. Pham & Rui J.P. deFigueiredo

Department of Electrical and Computer Engineering
Rice University, Houston, Texas 77251-1892

ABSTRACT

[illegible]

1. INTRODUCTION

In this paper we investigate the statistical properties of the maximum likelihood estimators of generalized p-Gaussian probability density function (gpG), and extend these results to the deconvolution problem under gpG white noise.

The gpG densities were initially introduced by Subbotin [illegible] in 1923 and consequently sometimes are called the Subbotin densities. In 1972, Miller and Thomas [illegible] used them as models for non-Gaussian noise in detection theory. Gray [illegible] in 1979 used the gpG as an earth model in his study on deconvolution of seismic signals. His approach is quite different from the one proposed by us. He assumed that the earth model (reflectivity sequence) is a random white gpG sequence and that the observed seismic trace consists of the source signal convolved with such an earth model and is contaminated by additive noise. He recovered (deconvolved) the reflectivity sequence by maximizing the generalized Varimax. Such an approach, even though based on physical grounds, is somewhat heuristic.

Almost all of the l_p deconvolution methods are based on *nonlinear programming techniques* [illegible],[illegible],[10] with the exception of l_1- and l_∞-deconvolution which have been executed using *linear programming techniques*[11].

In the following section, we define the gpG pdf, derive the maximum likelihood estimators (MLE's) of its mean and variance, and briefly summarize their properties. In section 3, we extend these developments to the general l_p deconvolution problem under gpG noise, that is to the determination of the system function of a linear system from the input-output data, the output being corrupted by additive gpG noise. Elsewhere[illegible], we have presented a solution to this problem based on a modification of the convex-simplex linear programming method. In the present paper, we show that this solution constitutes in fact an unbiased maximum likelihood estimate of the deconvolved system function, and obtain expressions for the Cramer-Rao lower bound and for an upper bound. For proofs and computer simulation, see the complete version which has been submitted to the trans. of ASSP.

2. CHARACTERIZATION AND PARAMETRIC ESTIMATION OF THE gpG pdf's

For p a positive integer, we define a gpG scalar random variable V as the one possessing a pdf of the form

$$[illegible] \tag{2.0.1}$$

where

$$[illegible] \tag{2.0.2}$$

$$[illegible] \tag{2.0.3}$$

and μ and σ^2 are the mean and variance of V. See figure 1.

We call gpG noise any sequence of iid random samples $V_1, V_2, \ldots$ from (2.0.1). It follows from (2.0.1) that the joint pdf for N such samples is

$$[illegible] \tag{2.0.4}$$

where $v = col(v_1, v_2, \ldots, v_N)$, and $V = col(V_1, V_2, \ldots, V_N)$. From this point on, we drop the subscripts indicating random variables when these are clear from the context. Also we have denoted the pdf as a likelihood function with μ and σ^2 as parameters.

2.1. Maximum Likelihood Estimates of the Mean and Variance.

The derivation of the estimate $\hat{\theta}_2$ of the variance σ^2 (assuming that μ is known) is straightforward. In fact, by setting

$$[illegible] = 0, \tag{2.1.1}$$

replacing (2.0.4) in (2.1.1) and solving for θ_2, we get

$$[illegible] \tag{2.1.2}$$

From (2.1.2) above, we can also derive a sequential algorithm for the maximum likelihood estimate of the variance of a gpG as follows.

Algorithm. Given $(\hat{\theta}_2)^{(N)}$ as the maximum likelihood estimate of the variance of a gpG, based on a set of N data points. Suppose we are given an additional data point, say v_{N+1}, the estimate of the variance based on the set of N+1 data points which can be computed using the previous estimate $(\hat{\theta}_2)^{(N)}$ and the new data point v_{N+1}

This work was supported by ONR contract N00014-85-K-0182

is:

$$(\hat{\sigma}_{ML}^{2})^{p/2} = \left[\frac{N(\ldots)^{p/2} + p\sum | v_{ML} - \mu |^{p}}{N+1}\right]^{2/p} \qquad (2.1.3)$$

For the estimate θ_1 of the mean μ the situation is slightly complicated for the case of odd p because μ in (2.1.2) resides within absolute value signs. For an even p, we set

$$\frac{\partial}{\partial\theta_1} \log [\, p(v) \mid \theta_1, \sigma^2 \,] = 0 \qquad (2.1.4)$$

which when combined with (2.0.4), after some manipulation leads to the following condition that θ_{1M} must satisfy

$$\sum_{k=0}^{p-1} C(p-1,k)\, a_k\, (-\theta_1)^{p-1-k} = 0 \qquad (2.1.5)$$

where

$$C(n,m) = \frac{n!}{m!\,(n-m)!} \qquad (2.1.6)$$

$$a_k = \sum_{i=1}^{N} (v_i)^k \qquad (2.1.7)$$

Above, $s = \mathrm{col}(a_0, a_1, \ldots, a_{p-1}) \in R^p$ is a sufficient statistic vector for θ_1.

The particular case in which $p=2$ corresponds to the Gaussian case and (2.1.5) can be inverted to give the familiar sample mean as the estimate of θ_1, i.e.

$$\hat{\theta}_1 = \hat{\mu}_{MLE} = \frac{a_1}{a_0} = \frac{1}{N}\sum_{i=1}^{N} v_i \qquad (2.1.8)$$

For p odd, maximization of (2.0.4) is equivalent to the minimization of

$$\sum_{i=1}^{N} | v_i - \theta_1 |^p = \sum_{i=1}^{N} (v_i - \theta_1)^p \,(\, \mathrm{sgn}(v_i - \theta_1) \,)^p \qquad (2.1.9)$$

The special cases correspond to $p=1$ and ∞ are of particular interest. In fact, it is well known that:

Proposition 2.1.1. For $p = 1$,

$$\hat{\theta}_1 = \hat{\mu}_{ML} = \mathrm{median}(v_1, v_2, \ldots, v_N) \qquad (2.1.10)$$

Proposition 2.1.2. The values of θ_1 which maximize the likelihood function are defined in the interval $[v_{max}-\sigma, v_{min}+\sigma]$ if this interval exists.

2.2. Properties. In this subsection, we summarize some of the important properties of the estimates derived above. Since $\sum_{i=1}^{N} | x_i - \theta_1 |^p$ is a strictly convex function of θ_1 for $1<p<\infty$, the minimizer $\hat{\mu}_{MLE}$ is unique. Furthermore, $\frac{\partial}{\partial\mu} f(v \mid \mu,\sigma)$ is absolutely integrable. Hence it follows (see [3]) that

Proposition 2.2.1. For $0<p<\infty$, $\hat{\mu}_{MLE}$ is consistent, asymptotically efficient, and asymptotically normal.

A stronger result can be obtained for the cases in which $p=1$ and 2. For $p=1$, the median and the mean are the same because $p(v|\mu,\sigma)$ is symmetric about the mean. So $\hat{\mu}_{MLE}$ is unbiased for $p=1$. As is also well known, the same holds for the case in which $p=2$ (Gaussian case). Hence it follows from the Cramer-Rao theorem that

Proposition 2.2.2. For $p=1$ and 2, $\hat{\mu}_{MLE}$ is , unbiased and hence the variance of the estimator is bounded below by the Cramer-Rao lower bound. The latter is given by $1/\alpha_\mu^2(\mu)$ where the Fisher information $\alpha_\mu^2(\mu)$ is given in (2.2.1).

A rather lengthy calculation leads to the following formula for Fisher's information $\alpha_\mu^2(\mu)$

Proposition 2.2.3. Let $(\partial^i/\partial\mu^i)\, f(v \mid \mu,\sigma)$ for $i=1,2$ exist and be absolutely integrable. Then the Fisher's information is

$$\alpha_\mu^2(\mu) = \frac{Np(p-1)\Gamma(\frac{1}{p})\Gamma(\frac{p-1}{p})}{\sigma^2\,[\,\Gamma(\frac{1}{p})\,]^2} \qquad (2.2.1)$$

Proposition 2.2.4. The maximum likelihood estimate $\hat{\mu}_{MLE}$ of the mean of the gpG is unbiased.

Having established the fact that $\hat{\mu}_{MLE}$ is unbiased, we obtain the Cramer-Rao lower bound for this estimate as the inverse of the Fisher's information:

Proposition 2.2.5. The Cramer-Rao lower bound for the maximum likelihood estimate $\hat{\mu}_{MLE}$ of the mean of the gpG is

$$E[(\hat{\mu}_{MLE} - \mu)^2] \geq \left[\frac{[\,\Gamma(\frac{1}{p})\,]^2}{Np(p-1)\Gamma(\frac{1}{p})\Gamma(\frac{p-1}{p})}\right]\sigma^2 \qquad (2.2.2)$$

The actual square error can be computed as follows

$$E[(\hat{\mu}_{MLE} - \mu)^2] = E[\hat{\mu}_{MLE}^2] - \mu^2 \qquad (2.2.3)$$

where $E[\hat{\mu}_{MLE}^2]$ is

$$E[\hat{\mu}_{MLE}^2] = J_l^T K J \qquad (2.2.4)$$

and J_l^T is the l_p generalized inverse of $[1\ 1\ \ldots\ 1]^T$, K the covariance matrix of v. Taking advantage of the following property of J_l^T:

$$\sum_i l_i = 1 \qquad (2.2.5)$$

it can be shown by direct multiplication that

$$E[\hat{\mu}^2] = \sigma^2 \sum_i l_i^2 + \mu^2 \qquad (2.2.6)$$

we obtain the result for proposition 2.2.6 as follows.

Proposition 2.2.6. The real square error of the maximum likelihood estimate of the mean μ of a gpG is:

$$E[(\hat{\mu}_{MLE} - \mu)^2] = \sigma^2 \sum_i l_i^2 \qquad (2.2.7)$$

From (2.2.5) and (2.2.7), we derive the following proposition on the upper bound of the estimate.

Proposition 2.2.7. The upper bound for the variance of the maximum likelihood estimate of the mean of a gpG as

$$E[(\hat{\mu}_{MLE} - \mu)^2] \leq \sigma^2 \qquad (2.2.8)$$

Turning now to $\hat{\sigma}_{MLE}^2$, it is clear from (2.1.2) that this estimate is unique for $1<p<\infty$. Hence

Proposition 2.2.8. The maximum likelihood estimate $\hat{\sigma}_{MLE}^2$ is biased.

Proof. We obtain the expected value of the estimate directly as:

$$E[\hat{\sigma}_{MLE}^2 \mid \sigma^2] = \left[\frac{p}{N}\right]^{2/p}\left[\frac{\Gamma((N+2)/p)}{\Gamma(N/p)}\right]\sigma^2 \qquad (2.2.9)$$

The result of proposition 2.2.8 gives the bias function for the maximum likelihood estimate of the variance as:

Proposition 2.2.9. The bias function for the ML estimate of the variance of the gpG is:

$$b(\hat{\sigma}_{MLE}^2 \mid \sigma^2) = \left[\left[\frac{p}{N}\right]^{2/p}\left[\frac{\Gamma((N+2)/p)}{\Gamma(N/p)}\right] - 1\right]\sigma^2 \qquad (2.2.10)$$

Fisher's information $\alpha_\sigma^2(\sigma^2)$ for this estimate is given by

$$\alpha_\sigma^2(\sigma^2) = \frac{Np}{4\sigma^4} \qquad (2.2.11)$$

From the Fisher's information and the bias function above, we calculate the Cramer-Rao lower bound for the biased maximum likelihood estimate of the variance of the gpG noise as

Proposition 2.2.10. The Cramer-Rao lower bound for the estimate of the variance of the gpG noise is:

$$E[(\hat{\sigma}^2_{ML}-\sigma^2)^2] \geq \left[\left[\frac{p}{N}\right]^{2/p}\left[1+\frac{p}{N}\right]\frac{\Gamma(\frac{N+4}{p})}{\Gamma(\frac{N}{p})}\right.$$

$$\left. -2\left[\frac{p}{N}\right]^{2/p}\frac{\Gamma(\frac{N+2}{p})}{\Gamma(\frac{N}{p})}+1\right]\sigma^4 \qquad (2.2.12)$$

with the special case of p = 2, the Cramer-Rao lower bound is

$$E[(\hat{\sigma}^2_{ML}-\sigma^2)^2] \geq \frac{2\sigma^4}{N} \qquad (2.2.13)$$

and for the special case of p = 1, the Cramer-Rao lower bound is

$$E[(\hat{\sigma}^2_{ML}-\sigma^2)^2] \geq \left[\frac{4N^2+[illegible]N+4}{N^2}\right]\sigma^4 \qquad (2.2.14)$$

Since the density function does not satisfy the efficient condition (except for the special case of p equals to 2), we calculate the real square error using multivariable calculus technique[4].

Proposition 2.2.11. The expected value of the square error of the maximum likelihood estimate of the variance of the gpG noise is:

$$E[(\hat{\sigma}^2_{ML}-\sigma^2)^2] = \left[\left[\frac{p}{N}\right]^{2/p}\frac{\Gamma(\frac{N+4}{p})}{\Gamma(\frac{N}{p})}\right.$$

$$\left. -2\left[\frac{p}{N}\right]^{2/p}\frac{\Gamma(\frac{N+2}{p})}{\Gamma(\frac{N}{p})}+1\right]\sigma^4 \qquad (2.2.15)$$

from which we obtain the results for the special case where p = 2 as

$$E[(\hat{\sigma}^2_{ML}-\sigma^2)^2] = \frac{2\sigma^4}{N} \qquad (2.2.16)$$

and for the special case where p = 1,

$$E[(\hat{\sigma}^2_{ML}-\sigma^2)^2] = \left[\frac{4N^2+[illegible]N+6}{N^2}\right]\sigma^4 \qquad (2.2.17)$$

Note that for the case of p = 2, which corresponds to the familiar Gaussian, the Cramer-Rao lower bound is equal to the expected square error, which confirms efficiency (this can be obtained by checking the efficient condition of the density function).

3. l_p DECONVOLUTION

The problem of l_p deconvolution can be modeled into a linear programming problem of the form proposed in [1]. Before defining the problem, we will present a mathematical definition of convolution of two discrete sequences $\{x_n|n=0,1,\ldots,N-1\}$ and $\{h_n|n=0,1,\ldots,M-1\}$ which is denoted by x*h as follows

$$x * h = \sum_{i=0}^{M-1} x(n-i)\,h(i) \qquad (3.0.1)$$

The problem of deconvolution is then stated as follows: Given the sequences $\{y_n|n=0,1,\ldots,M+N-2\}$ and $\{x_n|n=0,1,\ldots,N-1\}$. Find a sequence $\{h_n|n=0,1,\ldots,M-1\}$ such that $\{y_n|n=0,1,\ldots,M+N-2\}$ is the deconvolution of $\{x_n|n=0,1,\ldots,N-1\}$ and $\{h_n|n=0,1,\ldots,M-1\}$.

Consider the model of linear convolution as follows:

$$y = h * x + n \qquad (3.0.2)$$

where x is the input, h the transfer function of a linear system which can be expressed as Toeplitz matrix H operating on the vector x, n the additive zero-mean generalized p-Gaussian white noise, and y the output contaminated with noise n. Given the observed data y, the input x, we have an algorithm[1] that use linear programming technique to solve for a solution $\hat{h}$ which has the the following characteristics:

(i) $\hat{h}$ will give a minimized norm of the error function in an l_p normed space, i.e.

$$\| y - \hat{h} * x \|_p \leq \| y - h * x \|_p \qquad \text{for all } h \qquad (3.0.3)$$

(ii) If there is no additive noise, regardless of the space l_p normed space selected, $\hat{h}$ is unique, and results a zero error, i.e.

$$\| y - \hat{h} * x \|_p = 0 \qquad \text{for all } p \qquad (3.0.4)$$

3.1. Algorithm. Using the concept of linear programming technique, we derived an algorithm that solves for the optimal solution $\hat{h}$. For the derivation of the algorithm, see [1]. Below, we show this modified simplex algorithm to the l_p deconvolution problem.

(i) Initialization. Set the vector h to zero, i.e.

$$h_i^{(0)} = 0 \qquad i = 0,1,\ldots,M-1 \qquad (3.1.1)$$

Initialize an error vectors e and ε as

$$e_i^{(0)} = y_i \qquad (3.1.2a)$$

$$\varepsilon_i^{(0)} = | e_i^{(0)} | \qquad i = 0,1,\ldots,M+N-2 \qquad (3.1.2b)$$

(ii) Directional Search. Find the direction k^{th} that gives the smallest negative heuristic value θ_k:

$$\min_k \theta_k \qquad (3.1.3)$$

where

$$\theta_k = \sum_{i=0}^{M+N-1} [illegible]\, |x_{i-k}|\, [illegible] \qquad (3.1.4)$$

From this k^{th} direction, we assign the vector $d^{(n)}$ as

$$d_i^{(n)} = [illegible]\,| x_{i-k} | \qquad (3.1.5)$$

(iii) Stepsize Computation (or Line Search). Find a positive λ that optimizes an one-dimensional optimization problem:

$$\min_{[illegible]} \sum_{i=0}^{M+N-2} (\varepsilon_i^{(n)} + \lambda d_i^{(n)})^p \qquad (3.1.6)$$

where the constant λ_{max} is

$$\lambda_{max} = \min\left\{ -\frac{\varepsilon_i^{(n)}}{d_i^{(n)}} \right\} \qquad i \in \left\{ i \mid d_i^{(n)} < 0 \right\} \qquad (3.1.7)$$

then, the solution is updated as:

$$h_k^{(n+1)} = h_k^{(n)} + \lambda \qquad (3.1.8)$$

where k is the optimal direction in (ii), λ is the optimal solution of (3.1.6). The error vectors e and ε is updated as

$$e_i^{(n+1)} = e_i^{(n)} - \lambda x_i \qquad (3.1.9a)$$

$$\varepsilon_i^{(n+1)} = | e_i^{(n+1)} | \qquad i = 0,1,\ldots,M-1 \qquad (3.1.9b)$$

(iv) End Condition. If, in step (ii), no direction k^{th} would yield a negative heuristic value θ_k, then the solution is optimal. Otherwise, repeat steps (ii) through (iv).

The above algorithm evolves from the simplex and convex simplex algorithm, yet no tableau is constructed, thus saving a lot of buffer space and computation operations in the computer.

3.2. Statistical Properties. Denote the convolution process of a linear system by $T_x(h)$, where T_x is linear (this linearity can be verified easily by using the direct formula for convolution).

Theorem 3.2.1. Let X, Y be vector spaces, both real or both complex. Let T: D(T) → Y be a linear operator with domain D(T) ⊂ X and range R(T) ⊂ Y. Then if T^{-1} exists, it is a linear operator.

Proof: See Kreyszig[3], pp 88-89.

Unbiasedness. Using the above theorem, we can say that the deconvolution process is linear, i.e. T_x^{-1} is linear. Having this fact established, we can show that the estimate of h in equation (3.0.2) is unbiased as follows.

Theorem 3.2.2. The l_p deconvolution result in the presence of additive zero-mean gpG noise is unbiased.

Proof. Define the variable θ as

$$\theta = x * h \tag{3.2.1}$$

Then, the inverse operator T_x^{-1} will give

$$T_x^{-1}(\theta) = h \tag{3.2.2}$$

Let h_{MAE} be the estimate of h, i.e.

$$h_{MAE} = T_x^{-1}(y) \tag{3.2.3}$$

Then the expected value of h_{MAE} given h is

$$E[h_{MAE} | h] = E[T_x^{-1}(y) | h] \tag{3.2.4}$$

Since T_x^{-1} is linear, as shown previously, then

$$E[T_x^{-1}(y) | h] = T_x^{-1}(E[y | h]) \tag{3.2.5}$$

It is clear from (3.0.2) that

$$E[y | h] = \theta \tag{3.2.6}$$

Then

$$E[h_{MAE} | h] = T_x^{-1}(\theta) \tag{3.2.7}$$

Substitute equation (3.2.2) into (3.2.7), we have

$$E[h_{MAE} | h] = h \tag{3.2.8}$$

Therefore, the estimate h_{MAE} is unbiased. Q.E.D.

Cramer-Rao Bound. Under the condition of unbiased estimate, the Cramer-Rao lower bound is given in the form:

$$E[\sum_{i=1}^{N}(\hat{h}_i - h_i)^2] \geq \frac{N\sigma^2}{(p(p-1))\frac{\Gamma(\ldots)}{\Gamma(\ldots)}\Gamma(\frac{p-1}{p})\sum (x_m)^2} \tag{3.2.9}$$

which, for the special case of p = 2, the bound is

$$E[\sum_{i=1}^{N}(\hat{h}_i - h_i)^2] \geq \frac{N\sigma^2}{\sum (x_m)^2} \tag{3.2.10}$$

Now, with the efficient condition not satisfied, we calculate the actual error as follows

$$E[(\hat{h} - h)^T(\hat{h} - h)] = E[\hat{h}^T\hat{h}] - h^T h \tag{3.2.11}$$

with $\hat{h}$ being calculated from the generalized inverse T^T of X as

$$\hat{h} = T^T y \tag{3.2.12}$$

the actual error is

$$E[(\hat{h} - h)^T(\hat{h} - h)] = N\sigma^2 \operatorname{trace}(TT^T) \tag{3.2.13}$$

which gives the upper bound as

$$E[(\hat{h} - h)^T(\hat{h} - h)] \leq \frac{N\sigma^2}{\sum (x_i)^2} \tag{3.2.14}$$

For the special case of p = 2, the generalized inverse of X is the Penrose inverse, given as

$$X^{-1} = (X^T X)^{-1} X^T \tag{3.2.15}$$

which we use to compute the actual square error as

$$E[(\hat{h} - h)^T(\hat{h} - h)] = N\sigma^2 \operatorname{trace}(X^T X)^{-1} \tag{3.2.16}$$

4. CONCLUSION

In this paper we have investigated the properties of the gpG class of probability density functions with regard to the estimation of its parameters from a set of its N iid samples. This provided the setting for a statistical study of the solution of the l_p deconvolution problem obtained by the solution of an appropriate minimum norm problem in the l_p normed space. This solution we have shown to be unbiased and we have obtained for its variance the Cramer-Rao lower bound, and the upper bound.

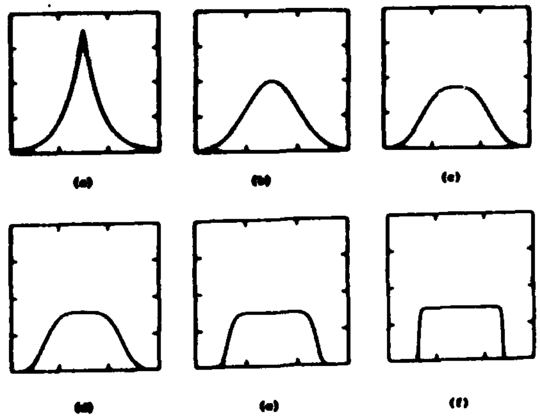


Figure 1. Density function for a generalized p-Gaussian r.v.

(a) p = 1 (c) p = 3 (e) p = 10

(b) p = 2 (d) p = 4 (f) p = 50

REFERENCES

[1] T. T. Pham, and R. J. P. deFigueiredo. A New Efficient Version of the Simplex Method for Application to a Class of Nonlinear Problems. Submitted to Mathematical Programming.

[2] S. S. Wilks. Mathematical Statistics. New York: John Wiley & Sons, Inc. (1962)

[3] E. Kreyszig. Introductory Functional Analysis with Applications. New York: John Wiley & Sons, Inc. (1978).

[4] J. E. Marsden, and A. J. Tromba. Vector Calculus. San Francisco: W. H. Freeman and Company. (1976).

[5] M. T. Subbotin. On the Law of Frequency of Errors. Matematicheskii Sbornik 31(1923), pp. 296-301.

[6] J. H. Miller, and J. B. Thomas. Detector for Discrete Time Signals in Non-Gaussian Noise. IEEE Trans. on Information Theory IT-18(1972), pp. 241-250.

[7] W. C. Gray. Variable Norm Deconvolution. Ph. D. thesis at Stanford University (1979).

[8] J. A. Fletcher, and M. D. Hebden. The Calculation of Linear Best l_p Approximation. Computer Journal 14 (1971), pp. 276-279.

[9] R. J. P. deFigueiredo, and T. A. Dwyer. Approximation Theoretic Methods for Nonlinear Deconvolution and Inversion. Information Science 31, pp. 209-220.

[10] P. L. Odell, T. G. Newman, and T. L. Boullion. Concerning Best Minimum p-Norm Estimators. Texas Center for Research Release.

[11] I. Barrodale, and A. Young. Algorithms for Best l_1 and l_∞ Linear Approximations of a Discrete Set. Numerische Mathematik 8(1966), pp. 295-306.

END

7-87

DTIC